

Sveučilište u Zagrebu

Filozofski fakultet

Odsjek za informacijske i komunikacijske znanosti

ak. god. 2015./2016.

Ivana Končić

Projekti masovne digitalizacije

završni rad

mentor: dr sc. Daniela Živković

Zagreb, 2016.

Sažetak

U uvodnome dijelu ovoga rada govori se o procesu digitalizacije knjižnične građe koja se pojavila kao odgovor na sve brži razvoj informacijske i komunikacijske tehnologije te je ubrzo postala poželjan faktor u svakoj knjižničnoj ustanovi. Osim što je omogućila bržu dostupnost informacija korisnicima knjižnica, uvelike je omogućila i dugo poželjno, trajno očuvanje knjižnične baštine za buduće generacije. Upravo su zbog toga mnoge svjetske knjižnice započele s digitalizacijom knjiga i ostale građe iz svojih zbirki. Do danas je velika količina knjiga već digitalizirana, no i dalje se ubrzano digitalizira u masovnim projektima uz uporabu napredne tehnologije te uz velika ekonomska ulaganja. Upravo je to glavna tema ovoga rada. Nakon iznošenja teorijskih postavki, slijedi analiza nekoliko većih projekata poput Projekta Gutenberg, Million Book, Open Content Alliance i Google Books. Budući da je masovna digitalizacija knjiga veliku pažnju javnosti počela privlačiti tek s razvojem Google Books-a, za detaljnije istraživanje njene problematike odabran je upravo taj projekt, kao i projekt Open Content Alliance kao primjer projekta koji teži vrlo sličnim ciljevima, a uz to je i decentraliziran, neprofitan, te na neki način više transparentan.

Ključne riječi

digitalizacija, masovni projekti digitalizacije, Gutenberg, Million Book, Open Content Alliance, Google Books

Sadržaj

Uvod	1
1. Digitalizacija	2
1.1. Digitalizacije knjige	2
2. Cilj i svrha digitalizacije	3
3. Projekti digitalizacije.....	4
4. Projekti masovne digitalizacije	7
4.1. Projekt Gutenberg	7
4.2. Projekt Million Book.....	9
4.3. Projekt Google Books	10
4.3.1. Proces digitalizacije i tehnologija koja se koristi	12
4.4. Projekt Open Content Alliance	16
4.4.1. Proces digitalizacije i tehnologija koja se koristi	17
5. Usporedna analiza projekata Open Content Alliance i Google Books	21
Zaključak	26
Literatura	28

Uvod

Kratku povijest elektroničke knjige moguće je pratiti od 1971. godine kada su volonteri Projekta Gutenberg započeli s digitalizacijom knjiga, čak prije nego što je široko rasprostranjen i dostupan bio Internet na kojem bi ih distribuirali. Prvi komercijalni paketi elektroničkih knjiga pojavili su se otprilike u isto vrijeme kao i prvi CD-ROM-ovi, budući da je tada bilo omogućeno skenirati cijeli tekst, prebaciti ga na računalo i pritom ga pretvoriti u digitalnu datoteku. U ranim devedesetima, e-knjige su postale sve više prihvaćene. Od tada, ideja čitanja knjiga u elektroničkom okruženju pridobila je široku prihvaćenost, zbog čega su knjižnice, ali i druge informacijske institucije sve češće u svoje zbirke počele uključivati upravo elektroničke knjige, kao i ostalu elektroničku građu. U novije vrijeme, kako oprema i rad osoblja potrebnog za provedbu opsežnih projekata digitalizacije poskupljuju, i kako tehnologija povećava kapacitet poslužitelja za pohranu ogromne količine digitalnih datoteka, formirao se određeni broj zajedničkih pothvata knjižnica i velikih korporacija i vlada kako bi zajednički podijelili troškove, ali i, takoreći, uživali u prednostima koje sa sobom donosi masovna digitalizacija.

Svijet razmjerno velikih i masovnih digitalizacijskih inicijativa još uvijek je relativno – malen. Postoji tek nekoliko projekata značajnije veličine, a još je manje onih koji teže univerzalnoj pokrivenosti i dostupnosti. U radu su istražena i opisana četiri projekta masovne digitalizacije knjiga: Projekt Gutenberg kao najstariji takav pothvat, Million Book kao primjer međunarodnog digitalizacijskog projekta koji ujedno potiče i provodi istraživanja digitalnih knjižnica, zatim Open Content Alliance kao konzorcij organizacija iz cijeloga svijeta koji su radili na stalno dostupnoj i javnoj arhivi digitalnih tekstova te Google Books kao primjer najvećeg projekta masovne digitalizacije – *istovremeno najambicioznijeg, ali i najspornijeg projekta ove vrste.*¹

Pomnije su istražena i uspoređena posljednja dva projekta – Open Content Alliance i Google Books – koji, iako naočigled u potpunosti različiti, dijele niz sličnosti. Ponekad je vrlo teško osvijestiti činjenicu da ne trebaju svi projekti s univerzalnim ambicijama izgledati kao najveći i najpoznatiji – kao što je to Google Books – i da, u stvari, te inicijative ne moraju zadržati svoju inicijalnu formu zauvijek. Upravo zbog toga, Open Content Alliance nudi pogled na

¹ Šapro-Ficović, M. Masovna digitalizacija knjiga : utjecaj na knjižnice. // Vjesnik bibliotekara Hrvatske. 54, 1/2(2011), str. 216-250. Dostupno na: <http://www.hkdrustvo.hr/vbh/broj/103> (15. kolovoza 2016.).

alternativni projekt koji teži prema vrlo sličnim ciljevima, a uz to je i decentraliziran, neprofitan, na neki način više transparentan.

1. Digitalizacija

U najširem smislu riječi, digitalizacija (eng. *digitalization*, od *digit*: znamenka) označava prevođenje analognoga signala u digitalni oblik. U užem smislu, ona se odnosi na pretvorbu teksta, slike, zvuka, pokretnih slika ili trodimenzionalnog oblika nekog objekta u digitalni oblik koji se može obrađivati, pohranjivati ili prenositi računalima i računalnim sustavima.² Digitalizacija se rabi u najrazličitijim područjima ljudskoga djelovanja – od profesionalnih do amaterskih primjena, iz čega se može zaključiti da se usporedno s razvojem Interneta i informacijske tehnologije, razvoj informacijskih znanosti proširio u većinu poslovnih i društvenih sfera. Sukladno tome, i potreba za digitalizacijom postala je nužna u gotovo svim ustanovama, uključujući i knjižnice.

1.1. Digitalizacije knjige

Digitalizacija knjige je proces pretvaranja fizičke knjige u digitalne slike, tj. elektroničke knjige putem skeniranja. Jednom kada je knjiga digitalno skenirana, slike su dostupne za brzu distribuciju, reprodukciju i čitanje na ekranu. Takve se slike spremaju u nekoliko vrsta formata kao što su Portable Document Format (PDF), Tagget Image File Format (TIFF) ili neki drugi format koji omogućava optičko prepoznavanje znakova (OCR) pretvarajući tako slike određenih stranica knjige u tekst i na taj način znatno smanjuje kapacitet za pohranu knjiga, ali i omogućava da se tekst ponovno oblikuje, pretraži itd.

Postupci digitalizacije, kao i uređaji kojima se ona obavlja, ovise o vrsti gradiva koje se digitalizira. Kada je riječ o knjigama, dokumentima i ostaloj građi čija je temeljna karakteristika slika ili tekst, digitalizacija se najčešće provodi skeniranjem. Vrijednija djela, poput povijesnih dokumenata i raritetnih knjiga digitaliziraju se snimanjem digitalnim fotografskim aparatima. Prema tome, digitalizacija slike u osnovi se sastoji od njena razlaganja na mrežu (raster) slikovnih elemenata; za svaki od tih elemenata binarno je kodiran njegov položaj u mreži te ton sive (kod jednotonske slike) ili tonovi osnovnih boja (kod slike u boji). Dobiveni digitalni zapis slike može biti bez sažimanja podataka (npr. datotečnoga formata TIFF) ili sa sažimanjem (JPEG), a takve se slike mogu naknadno računalno obraditi te ukloniti neželjene efekte.

² Digitalizacija. // Hrvatska enciklopedija. Mrežno izd. Zagreb : Leksikografski zavod Miroslav Krleža. Dostupno na: <http://www.enciklopedija.hr/natuknica.aspx?id=68025> (13. kolovoza 2016.).

Digitalizacija teksta provodi se istim postupkom kao i digitalizacija slike. Time se dobivaju digitalne slike teksta koje se mogu prikazati na zaslonu računala, čitati i prelistavati, ali ne i pretraživati i obrađivati. Da bi se to postiglo teksta se mora pretvoriti u računalno čitljiv oblik uz pomoć programa za optičko prepoznavanje znakova (eng. Optical Character Recognition, OCR), kojim se svakom znaku u tekstu dodjeljuje odgovarajući ASCII ili drugi binarni kod. Točnost kod suvremenih tiskanih tekstova je oko 99%, međutim, kod povijesnih tekstova i rukopisa ona se značajno smanjuje. Upravo zbog toga, osim postupaka skeniranja i fotografiranja, primjenjuje se i postupak prepisivanja. Ono je najdugotrajniji od nabrojanih postupaka i najčešće se primjenjuje kod rukopisa, tiskanog teksta s preslabim kontrastom ili tiskanog teksta kojem su pridodane rukom pisane bilješke.

2. Cilj i svrha digitalizacije

Općenito gledajući, „cilj i svrha digitalizacije najčešće su određeni namjerom ustanove da pomoću digitalnih preslika poboljša dostupnost građi, zaštiti izvornike, stvara nove proizvode i usluge te upotpuni fond ustanove. Osim toga, noviji trend predstavlja i provođenje postupka digitalizacije na zahtjev. Pri postavljanju ciljeva digitalizacije potrebno je voditi računa i o iskoristivosti računalnog medija, mogućnosti izgradnje osmišljenih digitalnih zbirki i privlačenju novih korisnika“.³ Budući da svaki od navedenih ciljeva digitalizacije postavlja određene zahtjeve, vrlo je važno da projekti digitalizacije na njih obrate pažnju, procjene njihovu razmjernu težinu te jasno definiraju kojoj će mjeri njihov krajnji proizvod odgovoriti na pojedini zahtjev.⁴

Dakle, razlozi zbog kojih se provode projekti digitalizacije, tj. temeljni ciljevi od kojih polaze, najčešće su sljedeći:⁵

1. *Digitalizacija radi zaštite izvornika* u početku je najčešći razlog. Postoje dva osnovna aspekta zbog kojih se provodi. Prvi je taj da se korisnicima na uporabu daju elektroničke verzije gradiva, a original se čuva od oštećenja, dok je drugi aspekt prezervacijski – što bi značilo da u slučaju gubitka ili uništenja originala elektronička kopija (u većoj mjeri) nadoknađuje gubitak.

³ Smjernice za odabir građe za digitalizaciju (radna verzija), 2007. // Nacionalni projekt „Hrvatska kulturna baština“, Ministarstvo kulture Republike Hrvatske. Dostupno na: https://bib.irb.hr/datoteka/590089.smjernice_odabir.pdf (14. kolovoza 2016.).

⁴ Stančić, H. Digitalizacija. Zagreb : Zavod za informacijske studije Odsjeka za informacijske znanosti Filozofskog fakulteta Sveučilišta u Zagrebu, 2009.

⁵ Ibid.

2. *Digitalizacija radi povećanja dostupnosti* provodi se iz razloga koji je jasan sam po sebi. Dok je original dostupan samo na jednom mjestu, dostupnost njegove digitalizirane, tj. elektroničke inačice gotovo je neograničena (tj. istovremeno je dostupna velikom broju korisnika).
3. *Digitalizacija radi stvaranja nove ponude i usluga* jedan je od najvažnijih razloga konverzije građe u digitalni oblik. Njome se naime otvara cijeli niz novih mogućnosti, kao što su razmjena metapodataka, ubrzano pretraživanje teksta, lakša analiza gradiva, virtualno spajanje fizički udaljenog gradiva i tome slično.
4. *Digitalizacija radi upotpunjavanja fonda i suradnje* odnosi se prije svega na nabavu digitalnih preslika građe iz drugih ustanova ili privatnih zbirki za koje se procjenjuje da su potrebne s obzirom na korisnike, svrhu i područje djelovanja ustanove.
5. *Digitalizacija na zahtjev korisnika* najbolje do izražaja dolazi kao dodatna usluga ustanove. Osim toga, može se iskoristiti kod postupaka određivanja prioriteta za digitalizaciju (kada je gradivo već odabrano), pri čemu se prednost može dati onim jedinicama građe koju korisnici najprije zatraže.⁶

Osim navedenih razloga, neki projekti ciljaju i na druge učinke kao što su promidžba, senzibiliziranje zajednice za djelatnost ustanove, jačanje institucionalnih sposobnosti, pa i zarada.⁷ Nadalje, ne samo što omogućuje dostupnost literature većem broju korisnika, već i pridonosi očuvanje knjižnične baštine stare i po nekoliko stotina godina.

3. Projekti digitalizacije

Iako pogrešno, ustaljeno je shvaćanje koje digitalizaciju i projekte digitalizacije podrazumijeva jednakima. Međutim, digitalizacijski projekti puno su više od samo digitalizacije: oni uključuju mnogo drugih pitanja, aktivnosti i problema. To su, na primjer, odabir i priprema zbirke građe koja će se digitalizirati, zatim rješavanje pitanja autorskih

⁶ Smjernice za odabir građe za digitalizaciju (radna verzija), 2007. // Nacionalni projekt „Hrvatska kulturna baština“, Ministarstvo kulture Republike Hrvatske. Dostupno na: https://bib.irb.hr/datoteka/590089.smjernice_odabir.pdf (14. kolovoza 2016.).

⁷ Neomedia Komunikacije. Priprema i planiranje projekta digitalizacije. 2008. // Hrvatska kulturna baština. Dostupno na: <http://www.kultura.hr/Sudjelujte/Digitalizacija/Priprema-i-planiranje-projekta-digitalizacije> (17. kolovoza 2016.).

prava, stvaranje metapodataka, implementacija strategija očuvanja i zaštite digitalnih slika itd.⁸

Okvir za izradu i upravljanje arhivom digitalnih slika, tj. arhivom u koji će se pohranjivati proizvodi projekta digitalizacije, može se podijeliti u tri glavna područja. To su:

1. Snimanje i stvaranje podataka (*Data Capture and Creation*) – dokument se digitalizira, dodaju se metapodaci koji opisuju sliku (dokument), te se podaci (=digitalni zapis: slika i povezani metapodaci) pohranjuju ili arhiviraju za buduću uporabu
2. Pristup i dostava podataka (*Data Access and Delivery*) – potrebno je implementirati mehanizam dostave (preuzimanja) podataka koji će korisnicima osigurati pristup arhivi
3. Upravljanje digitalnom zbirkom (*Managing the Digital Collection*) – proizvodnja visokokvalitetnih digitalnih zbirka slika koja se pridržava standarda i dobre prakse, bit će osigurana koordinacijom prvog i drugog područja

Svako područje ima vlastiti skup ključnih aktivnosti i dok se mnoge od njih provode linearnim slijedom, one nisu izričito odvojene već tvore kompleksnu strukturu međusobno povezanih zadataka. Ključne aktivnosti tri navedena područja, prikazane su u tablici.⁹

Snimanje i stvaranje podataka	Pristup podacima i dostava	Upravljanje digitalnom zbirkom
rukovanje građom i priprema za skeniranje	pretraživanje i dohvaćanje	stvaranje baze podataka, dizajn sustava
skeniranje građe (hardver i softver)	upravljanje pristupom; rad s korisnicima	upravljanje procedurama i tijekom rada
sažimanje datoteka u željeni format		osiguranje kvalitete
autorska prava, prava intelektualnog vlasništva, etika i zaštita podataka		upravljanje projektom
metapodaci: opis slike, indeksiranje i katalogizacija		čuvanje i skladištenje digitalnih datoteka

⁸ Youngs, K. Managing the digitisation of library, archive and museum materials. London: NPO, 2001. Dostupno na: <http://www.collectionstrust.org.uk/media/documents/c1/a105/f6/000505.pdf> (14. kolovoza 2016.).

⁹ prema tablici preuzetoj iz rada autorice K. Youngs (2001).

Važno je naglasiti da ne postoji točno određen način na koji se provode projekti digitalizacije. Svaki je projekt jedinstven zbog svoje definirane svrhe (tj. onoga za što će se digitalizirani dokumenti koristiti, ili pak zbog razloga iz kojih se digitalizacija provodi – npr. zaštita, dostupnost itd.), kao i zbog razlike u građi koja će se digitalizirati te stanja u kojem se nalazi (što dalje utječe na tehnike i tehnologije skeniranja). Sve tehničke odluke temelje se upravo na tim faktorima. Međutim, korištenje okvira koji uključuje smjernice za dobru praksu omogućava projektima održavanje i napredovanje kroz bilokakve tehnološke promjene.

Općenito gledajući, projekte digitalizacije možemo razlikovati po mnoštvu različitih kriterija. Na skali digitalizacije, Coyle projekte dijeli na nemasovne, razmjerno velike i masovne projekte.¹⁰

Projekti nemasovne digitalizacije temelje se na pažljivom i individualnom odabiru građe koja će se digitalizirati. To je oblik digitalizacije koji proizlazi iz projekata zaštite građe, čija je svrha stvaranje zamjenske kopije dokumenta koji se nalazi u procesu raspadanja ili pak omogućavanje široke dostupnosti i lakšeg pristupa rijetkim zbirkama građe. Primjer takve vrste pothvata je digitalizirana građa iz Zbirke starih knjiga i rukopisa Nacionalne i sveučilišne knjižnice u Zagrebu koja nudi pogled u najvrjedniji dio svog fonda, te donosi izbor značajnih djela hrvatske i europske pisane baštine.

Druga vrsta projekata odnosi se na razmjerno velike, tj. *large-scale* projekte digitalizacije. Unatoč tome što i njihov rezultat može biti veliki broj digitalizirane građe (tj. skeniranih stranica), za razliku od masovnih projekata, oni su selektivni, budući da ih zanima stvaranje određene zbirke ili niza dokumenata po specifičnim kriterijima. Jedan od primjera takvih projekata je American Memory – projekt Kongresne knjižnice u SAD-u koji okuplja digitaliziranu građu o američkoj povijesti: dokumente, fotografije, zemljovide i mnoge druge oblike građe.

I zadnja vrsta projekata koje navodi Coyle, odnosi se na projekte masovne digitalizacije. Iako nema univerzalno prihvaćenu definiciju, dvije su karakteristike koje se s njom najčešće povezuju. Prva se odnosi na relativni nedostatak selektivnosti u odnosu na manje projekte, a druga na veliku brzinu procesa u smislu digitalne konverzije dokumenta, kao i stvaranja metapodataka. Osim toga, tu su i veliki obujam digitalizirane građe te skoro potpuna

¹⁰ Coyle, K. Mass digitization of books. // Journal of Academic Librarianship 32, 6(2006). Dostupno na: <http://www.kcoyle.net/jal-32-6.html> (15. kolovoza 2016.).

automatizacija procesa. Prema Coyle (2006) masovna digitalizacija prije svega podrazumijeva konverziju građe u industrijskim razmjerima, tj. konverziju cjelovitih fondova knjižnica bez specifičnog odabira ili selekcije. Prema tome, cilj masovne digitalizacije nije stvaranje posebnih zbirki, već digitalizacija *svega*, u ovome slučaju – svake knjige ikad tiskane.

4. Projekti masovne digitalizacije

Iz mnogih digitalizacijskih projekata aktivnih 1980-ih i 1990-ih, proizišao je koncept digitalne knjižnice – uređene zbirke digitalnih sadržaja stvorene zbog specifičnih razloga i integrirane s knjižničnim katalogom te fizičkom zbirkom.¹¹ Općenito gledajući, knjižnice digitaliziraju građu kako bi učinile boljim ono što su oduvijek radile – povezivale sadržaj s korisnicima. U tradicionalnom modelu knjižnice, korisnici ju posjećuju kako bi posudili knjige, časopise i ostalu građu, dok je u digitalnoj knjižnici sadržaj gotovo odmah dostupan za korištenje. Sukladno tome, utopijska vizija univerzalne knjižnice polako postaje stvarnost. Tom su se vizijom početkom ovoga tisućljeća, u svojim začecima vodili i neki od najvećih projekata razmjerno opsežne i masovne digitalizacije. Projekti poput Gutenberga, Million Booka (poznat i pod nazivom the Universal Library), Google Booksa (nastao pod nazivom Google Print) te Open Content Alliancea samo su neki od primjera čija je vizija stvaranje opsežne zbirke digitalnih dokumenata i ostalih digitalnih sadržaja, a čiji je motiv na prvom mjestu omogućiti i poboljšati pristup toj građi i informacijama.

4.1. Projekt Gutenberg

Početak masovne digitalizacije, a ujedno i pojavu e-knjige, moguće je povezati s 1971. godinom kada je Michael Hart, tadašnji student Sveučilišta u Illinoisu, sveučilišnom mrežom počeo distribuirati knjige u digitalnoj formi. Uvidjevši prednosti raspačavanja djela putem mreže te mogućnost da se digitalizirani tekst čita na zaslonu računala, pokrenuo je projekt pod nazivom Gutenberg. Stvaranjem prvoga elektroničkog dokumenta ručnim unošenjem teksta *Deklaracije o neovisnosti SAD-a*, korištenjem nove tehnologije (koju Hart naziva tehnologijom repliciranja) kao i utemeljenjem filozofske osnove za stvaranje internetske zbirke elektroničkih knjiga,¹² Projekt Gutenberg postao je simbolom digitalizacije i svjetske digitalne knjižnice.

¹¹ Deegan, M., Sutherland, K. Transferred illusions : digital technology and the forms of print. Farnham: Ashgate, 2009.

¹² Hart, M. The history and philosophy of project Gutenberg. 1992. // Project Gutenberg. Dostupno na: http://www.gutenberg.org/wiki/Gutenberg:The_History_and_Philosophy_of_Project_Gutenberg_by_Michael_Hart (15. kolovoza 2016.).

Do 1989. godine tekstovi su se unosili ručno i to u vrlo običnom formatu zvanom Plain Vanilla ASCII. Nakon te godine počeli su se koristiti skeneri i OCR, te je projekt počeo poprimati današnji oblik. Osim u formatima za računalo, knjige su danas dostupne i u raznim drugim formatima pogodnima za iPhone, Kindle, te različite druge čitače.¹³

Što se tiče odabira građe za digitaliziranje, u projektu se digitaliziraju knjige kojima je isteklo autorsko pravo, a osim toga uključuju se i djela zaštićena autorskim pravom za koje su dobili odobrenje autora. Međutim, broj takvih digitaliziranih djela veoma je malen. Statistika digitalizacije odražava polagani razvoj: objavljivanjem djela King James Bible 1989. godine zbirci se dodaje tek deseti naslov, stoti naslov 1994. godine, objavljivanjem svih djela Williama Shakespearea, zatim tisućiti naslov 1997. godine objavljivanjem Danteovog Pakla, a 2003. godine objavljivanjem Magna Carte, Projekt dobiva deset tisućiti naslov. Danas Projekt Gutenberg za slobodno preuzimanje nudi više od 52 tisuće besplatnih elektroničkih knjiga različitih formata, a zajedno sa srodnim grupacijama i partnerima nudi više od 100 tisuća elektroničkih knjiga koje se mogu preuzeti ili čitati online.¹⁴

Glavno je poslanje projekta poticati stvaranje i raspačavanje elektroničkih knjiga, a pri tome poticati i pomagati sve one koji su zainteresirani za rad na ovome projektu.¹⁵ Naime, svatko se može priključiti i doprinijeti radu na projektu, bilo da je riječ o izradi elektroničke knjige ili lektoriranju. Digitalizaciju i korekturu provodi tisuće volontera iz cijelog svijeta o kojima ovisi i koje knjige i na kojem jeziku će se objavljivati. Najviše je proznih djela na engleskom jeziku, upravo zato jer je najviše volontera koji to žele digitalizirati.

Kritike na račun projekta najviše su se odnosile na stručne nedostatke elektroničkih inačica djela kao što su izostavljanje informacija o korištenom izdanju knjige i tome slično. Također, dosta je prigovora bilo i zbog prevelike zastupljenosti sabranih djela određenih pisaca (npr. Charles Dickens), dok su druga manje zastupljena.¹⁶ Pokazalo se ipak da projekt vodi računa o kritikama i ispravlja greške kad se na njih upozori, što je možda i čimbenik njihovog dugogodišnjeg uspješnog poslovanja i postojanja.

¹³ Šapro-Ficović, M. Masovna digitalizacija knjiga : utjecaj na knjižnice. // Vjesnik bibliotekara Hrvatske. 54, 1/2(2011), str. 216-250. Dostupno na: <http://www.hkdrustvo.hr/vbh/broj/103> (15. kolovoza 2016.).

¹⁴ Free ebooks by project Gutenberg. // Project Gutenberg. Dostupno na: http://www.gutenberg.org/wiki/Main_Page (15. kolovoza 2016.).

¹⁵ Hart, M. Project Gutenberg mission statement. 2007. // Project Gutenberg. Dostupno na: http://www.gutenberg.org/wiki/Gutenberg:Project_Gutenberg_Mission_Statement_by_Michael_Hart (15. kolovoza 2016.).

¹⁶ Šapro-Ficović, M. Masovna digitalizacija knjiga : utjecaj na knjižnice. // Vjesnik bibliotekara Hrvatske. 54, 1/2(2011), str. 216-250. Dostupno na: <http://www.hkdrustvo.hr/vbh/broj/103> (15. kolovoza 2016.).

4.2. Projekt Million Book

Projekt Million Book osnovan je 2002. godine na Sveučilištu Carnegie Mellon u suradnji Nacionalne zaklade za znanost SAD-a s brojnim sveučilištima i radnim centrima u Kini i Indiji, a nešto kasnije i u Egiptu. Ono što ga možda najviše izdvaja od drugih je upravo činjenica da je to izrazito međunarodan projekt koji, kroz nekoliko servera u SAD-u, Kini i Indiji, pristup digitaliziranoj građi omogućuje u širim zemljopisnim razmjerima, čime se dodatno čuva cjelovitost zbirke.¹⁷ Od 26 institucija koje su surađivale na projektu, neke su doprinosile stvaranju sadržaja, dok su ostale radile na istraživanjima digitalnih knjižnica.

Specifična karakteristika ovog projekta upravo je u njegovom opsežnom programu istraživanja digitalnih knjižnica. Program uključuje proučavanje skladištenja velikog broja informacija i njihovo upravljanje, zatim tehnike skeniranja i obrada slika, optičko prepoznavanje znakova i obrada teksta, indeksiranje itd. Osim toga, bave se i problematikom autorskih prava, upravljanjem digitalnim pravima i sličnim temama.¹⁸

Sam proces skeniranja započeo je 2002. godine u nekoliko središta u Indiji i Kini, a završio u srpnju 2007. godine istekom izvornih fondova. Od početka skeniranja, projekt je premašio svoj zadani cilj od 1 milijuna knjiga, te je sveukupno skenirano oko 1,5 milijuna knjiga – dnevno se skeniralo oko 7 tisuća knjiga na čemu je radilo oko tisuću ljudi. Projekt je obuhvatio knjige objavljene na 23 jezika, velika većina ih je na kineskom (970 tisuća) i engleskom (360 tisuća), zatim 50 tisuća na južnoindijskom telugu, te 40 tisuća na arapskom jeziku.¹⁹

Iako je krajem 2007. godine projekt očigledno stao s tempom digitalizacije koju je do tada provodio, partneri na projektu i dalje su nastavili raditi zajedno u pojedinačnim istraživačkim radovima po svijetu, što je dokaz svestranog i globalnog utjecaja projekta. Svojom su vizijom sveobuhvatnog, slobodnog pristupa knjigama kao svjetske baštine, kao i s mnogim tehnološkim i praktičnim inovacijama u području digitalizacije knjige, služili kao poticaj mnogim drugim projektima i pritom utjecali na promjenu općeg mišljenja o reprezentaciji informacija sadržanih u knjigama.

¹⁷ Šapro-Ficović, M. Masovna digitalizacija knjiga : utjecaj na knjižnice. // Vjesnik bibliotekara Hrvatske. 54, 1/2(2011), str. 216-250. Dostupno na: <http://www.hkdrustvo.hr/vbh/broj/103> (15. kolovoza 2016.).

¹⁸ Rieger, O. Preservation in the age of large-scale digitization. Washington, D.C.: Council on Library and Information Resources, 2008. Dostupno na: <https://www.clir.org/pubs/reports/pub141/pub141.pdf> (16. kolovoza 2016.).

¹⁹ Online library gives readers access to 1.5 million books. 2007. // Carnegie Mellon University. Dostupno na: https://www.cmu.edu/news/archive/2007/November/nov27_ulib.shtml (15. kolovoza 2016.).

4.3. Projekt Google Books

Masovna digitalizacija knjiga usko je povezana s Googleovom prošlosti. Njegovi osnivači Sergey Brin i Larry Page, kao studenti Standfordskog sveučilišta 1996. godine, razvijali su razne tehnologije i bavili se konceptom digitalnih knjižnica. Njihov rad bio je usmjeren na proučavanje primjene indeksiranja na sadržaj digitaliziranih knjiga. Rezultat tog projekta, odnosno tehnologiju koja je nastala, primijenili su na arhive web stranica i ostalo je, kako se kaže, povijest – razvili su tehnologiju pod nazivom PageRank, karakterističnu za Google.²⁰ Međutim, interes za opsežnom digitalizacijom nije nestao. Godine 2002. započeli su s planiranjem faza vlastitog projekta u početku poznatog pod nazivom Google Print. Nakon što su razvili vlastitu metodu skeniranja i riješili sve tehničke probleme, 2004. godine na Frankfurtском sajmu, Google je objavio vijest o pokretanju projekta Google Print čija će primarna svrha biti *olakšati ljudima pronalazak značajnih knjiga*.²¹

U projekt je isprva bilo uključeno pet velikih knjižnica: četiri sveučilišne – Harvard, Stanford, Michigan i Oxford te New York Public Library. Do 2005. godine projektu su se pridružili suradnici iz osam europskih zemalja: Austrije, Belgije, Francuske, Njemačke, Italije, Nizozemske, Španjolske i Švicarske, a dalje je broj suradnika bio u konstantnom porastu.

Svoju zbirku digitaliziranih knjiga, Google stvara na dva načina: kroz *Google Books Publisher* (tzv. Partnerski program, za izdavače i autore) i *Google Books Library* (Knjižnični program).

Partnerski program marketinški je program za nakladnike i autore koji prije svega služi za promicanje njihovih djela. Oni se mogu registrirati u Partnerski program, pri čemu im je omogućeno besplatno uključivanje vlastitih knjiga u Googleov indeks pretraživanja te primanje polovice prihoda od oglasa i reklama koje Google uparuje s rezultatima pretraživanja. Pri pretraživanju knjiga u Partnerskom programu, pojavljuje se bibliografski zapis, broj stranice na kojoj se nalazi traženi pojam te do dvije stranice s jedne i druge strane ključne riječi. Uz rezultate pretraživanja, prikazuju se i linkovi na web stranice koje prodaju traženu knjigu (npr. Amazon), uključujući i link na web stranicu samog izdavača. Iako Google skenira i pohranjuje cijeli tekst knjige na svojim serverima, nekoliko stranica

²⁰ Horvat, A., Živković, D. Između javnosti i privatnosti : knjižnice u vremenu e-knjige. Zagreb : Hrvatska sveučilišna naklada, 2012.

²¹ About Google Books. // Google Books. Dostupno na: <https://books.google.com/googlebooks/about/> (16. kolovoza 2016.).

namjerno se ne prikazuje. Uz to je korisnicima onemogućeno ispisivanje i kopiranje slika, tj. dokumenata.

Dok u Partnerskom programu izdavači i autori biraju koje će se knjige skenirati, u Knjižničnom programu Google skenira zbirke velikih sveučilišnih i narodnih knjižnica bez obzira na stanje autorskih prava. Fond tipične sveučilišne knjižnice koji Google skenira sadrži 25% knjiga izdanih prije 1923. godine. Budući da one nisu zaštićene autorskim pravima, njihove elektroničke inačice mogu se pročitati i preuzeti u cijelosti. No za 75% knjiga izdanih nakon te godine smatra se da su još uvijek pod zaštitom autorskih prava. Njihove elektroničke inačice ne mogu se u cijelosti čitati, već je dopušten pristup samo 20% sadržaja u obliku kraćih izvadaka rečenica iz teksta (eng. snippets).²² Obično se stavljaju po 3 izvatka kao dodatak bibliografskim podacima o knjizi. U zamjenu za pristup knjigama, Google omogućuje osoblje i financijsku potporu. Pritom stvara dvije digitalne kopije dokumenta, od kojih jednu pohranjuje na svojim serverima (gdje ju ima obvezu čuvati 20 godina), a drugu šalje u knjižnicu iz koje je dokument potekao.²³

Općenito gledajući Partenrski program Google Books-a dobio je uglavnom pozitivne reakcije izdavača, autora i korisnika. Nasuprot njemu, Knjižnični program meta je mnogih kritika. Izdavači su se žalili na činjenicu da im projekt, takoreći, oduzima prihode koje bi inače zaradili u tom području poslovanja. Budući da knjižnice sada mogu dobiti digitalizirane kopije svojih tiskanih knjiga od Google-a, one više neće kupovati te iste e-knjige izravno od izdavača. Kao odgovor na te kritike, Google je program opisao kao *elektronički kartični katalog* koji korisnicima pomaže u lociranju i pronalaženju informacija. Google tvrdi da program pogoduje nositeljima autorskog prava, kao i izdavačima, jer knjige čini dostupnijima, a samim time im povećava vjerojatnost da će se kupiti.²⁴

Do sada, mnoge su istaknute knjižnice prihvatile Googleovu ponudu za suradnju, uključujući New York Public Library, knjižnice sveučilišta Michigan, Harvard, Stanford i Oxford, zatim Austrijska nacionalna knjižnica, Bavarska državna knjižnica itd. Međutim, otpor knjižnica poput Boston Public Library i Smithsonian Institution sugerira da mnoge akademske i neprofitne institucije imaju namjeru slijediti viziju Interneta kao globalnog repozitorija znanja oslobođenog poslovnih interesa i ograničenja. Iako bi ovaj projekt milijune knjiga moga

²² Horvat, A., Živković, D. Između javnosti i privatnosti : knjižnice u vremenu e-knjige. Zagreb : Hrvatska sveučilišna naklada, 2012.

²³ Ibid.

²⁴ Balas, J. L. By digitizing, are we trading future accessibility for current availability? // Computers in Libraries 27, 3(2007), str. 30-32. Dostupno na: <http://connection.ebscohost.com/c/editorials/24353509/by-digitizing-are-we-trading-future-accessibility-current-availability> (13. kolovoza 2016.).

učiniti dostupnima milijunima korisnika diljem svijeta, neke knjižnice, kao i znanstvenici i istraživači pokazuju određenu zabrinutost. Naime, ukoliko bi došlo do toga da jedna institucija/tvrtka počne dominirati digitalizacijom građe, tu bi dominaciju mogla iskoristiti za vlastitu komercijalnu dobit.²⁵

Nadalje, s najžešćim se kritikama Google Books susreo na području tumačenja i primjene zakona o autorskim pravima. Godine 2005., manje od godine dana nakon što je objavljena namjera da Google započinje sa skeniranjem knjiga, udruge Authors Guild i American Association of Publishers podnijele su odvojene tužbe u kojima navode da Google krši Zakon o autorskim pravima time što digitalizira i nudi na korištenje građu zaštićenu autorskim pravima, a sve kako bi ostvario financijsku dobit.²⁶ Prilikom pokušaja dogovora oko kompromisa s članovima udruga, Google je dobrovoljno pristao zaustaviti skeniranje zaštićene građe, a Američko udruženje izdavača predložilo je rješenje temeljeno na jedinstvenom ISBN broju koji bi se dodijelio svakoj knjizi objavljenoj nakon 1967. godine. Koristeći ISBN brojeve, Google bi mogao utvrditi koja su djela zaštićena autorskim pravima te potom kontaktirati izdavače i autore kako bi dobio dopuštenje za skeniranje. Međutim, taj je prijedlog Google odbio, pregovori su propali, a skeniranje zaštićenih knjiga se nastavilo. Ipak, kao odgovor na sve veće pritiske, Google je izdavačima i vlasnicima autorskih prava ipak omogućio opciju isključenja iz sudjelovanja u projektu (tzv. opt-out).²⁷

4.3.1. Proces digitalizacije i tehnologija koja se koristi

Godine 2002. kada su u Google-u tek započeli eksperimentirati s digitaliziranjem, za skeniranje knjige od 300 stranica bilo je potrebno otprilike 40 minuta. Danas, u samo sat vremena skener može digitalizirati oko 6000 stranica. Do 2015. godine skenirano je više od 25 milijuna svezaka, uključujući i tekstove na 400 jezika iz više od 100 zemalja.²⁸

Proces digitalizacije (ilustriran na Slici 1.) započinje s identifikacijom zbirke, i to tako da Google i knjižnica zajedno utvrde koje su jedinice građe, od sveukupnog fonda, prikladne za

²⁵ Hafner, K. Libraries rebuff Google and Microsoft on offers to place books on web. 22. listopada 2007. // The New York Times Magazine. Dostupno na: <http://www.nytimes.com/2007/10/22/technology/22iht-library.1.7992196.html> (15. kolovoza 2016.).

²⁶ Balas, J. L. By digitizing, are we trading future accessibility for current availability? // Computers in Libraries 27, 3(2007), str. 30-32. Dostupno na: <http://connection.ebscohost.com/c/editorials/24353509/by-digitizing-are-we-trading-future-accessibility-current-availability> (13. kolovoza 2016.).

²⁷ Dye, J. The digital rights issues : behind book digitization projects. // EContent 29, 1(2006). Dostupno na: <http://www.econtentmag.com/Articles/Editorial/Feature/Scanning-the-Stacks-The-Digital-Rights-Issues-Behind-Book-Digitization-Projects-14884.htm> (14. kolovoza 2016.).

²⁸ Heyman, S. Google Books : a complex and controversial experiment. 2015. // The New York Times Magazine. Dostupno na: http://www.nytimes.com/2015/10/29/arts/international/google-books-a-complex-and-controversial-experiment.html?_r=0 (17. kolovoza 2016.).

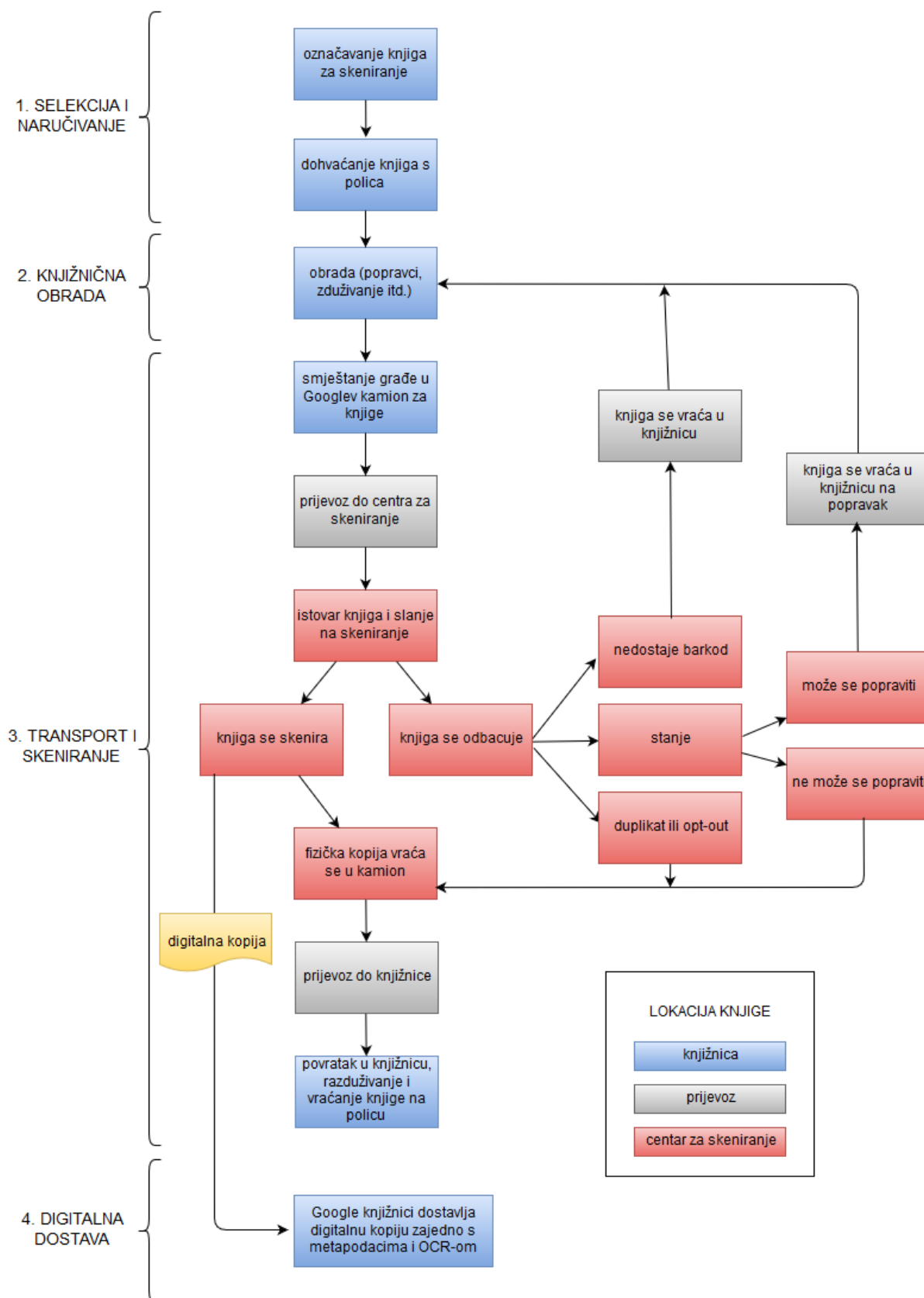
digitalizaciju – uzimajući u obzir materijal, veličinu knjige, njenu vrijednost i druge čimbenike. Nakon što je zbirka za digitalizaciju odabrana, knjižnično osoblje priprema knjige, skenira njihove barkodove i *posuđuje* ih Googleu za vrijeme trajanja procesa. Nakon toga, građa se transportira do objekta u kojem Googleovo osoblje obavlja digitalizaciju. Ukoliko je riječ o sveučilišnoj knjižnici, objekt se najčešće nalazi unutar kampusa. Općenito, svi objekti u kojima se obavlja digitalizacija, odvojeni su od knjižnice čija se građa skenira.²⁹ Centar u kojem se skeniranje obavlja sastoji se od brojnih postaja od kojih je svaka opremljena specijaliziranom Googleovom tehnologijom skeniranja. Za svakom postajom nalazi se po jedna osoba koja ručno okreće stranice dok kamere iznad knjige fotografiraju obje stranice. Neposredan proizvod tog procesa skup je slika preuzetih s digitalne kamere izravno na računalo, koje se nakon toga šalju Googleov server zbog daljnje obrade. Nakon što su stvorene master verzije fotografija, one se komprimiraju i pakiraju zajedno u jednu PDF datoteku koja je spremna za preuzimanje. Nakon što Google dovrši svaku seriju knjiga, one se vraćaju knjižničnom osoblju koji ih razdužuje i vraća natrag na police. Također, osoblje dodatno pregledava određeni uzorak digitaliziranih datoteka koje im Google dostavlja kako bi osigurali da zadovoljavaju potrebne kriterije.³⁰

Da bi knjižnice mogle preuzeti digitalne datoteke nastale skeniranjem njihove građe, Google je osigurao poseban mehanizam pristupa. Uz to im daje i dozvolu za distribuciju tih datoteka za vlastite usluge bez nekih značajnijih ograničenja. Međutim, komercijalna i masovna distribucija je, naravno, zabranjena.

Ostalim korisnicima, digitalizirana građa dostupna je na posebnoj web-stranici <http://books.google.com>, a osim toga, knjigama se može direktno pristupiti i pretraživanjem u standardnoj Googleovoj tražilici.

²⁹ Said, C. Revolutionary chapter : Google's ambitious book-scanning plan seen as key shift in paper-based culture. 20. prosinca 2004. // San Francisco Chronicle. Dostupno na: <http://www.sfgate.com/business/article/Revolutionary-chapter-Google-s-ambitious-2662491.php> (13. kolovoza 2016.).

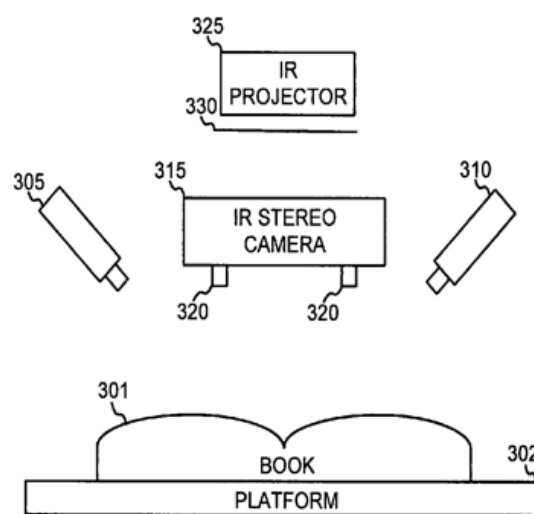
³⁰ Karle-Zenith, A. Google Book Search and the University of Michigan. 20. veljače 2007. Dostupno na: http://eprints.rclis.org/9011/1/Zenith_Google.pdf (15. kolovoza 2016.).



Slika 1. Google Books - tijek procesa digitalizacije knjige (prema Jones, 2014)

Iako je Google u samim počecima projekta bio vrlo tajanstven po pitanju tehnologije koju koristi prilikom skeniranja, 2009. godine ipak je objavljen Patent inovativnog sustava koji se koristi. U osnovi, sustav koristi višestruke kamere i trodimenzionalno modeliranje kako bi identificirao uvez knjige i ispravio prirodnu zakrivljenost stranica knjige koja leži ravno na podlozi. Kada se otvorena knjiga *normalno* fotografira odozgo, rubovi stranice prikazat će se kao zakrivljene linije. Ovom se metodom to ispravlja – izradom 3D modela svake stranice, sustav proizvodi ravne stranice bez potrebe da svaku fizički poravna što bi zahtijevalo uporabu neke od destruktivnih metoda poput uništavanja uveza ili staklenih ploča za individualno ravnjanje stranica.³¹ Osim što se takvim metodama nanosi šteta originalnoj građi, neučinkovito je za masovnu digitalizaciju.

Skica sustava prikazana je na Slici 2. na kojoj se vidi pozicioniranje knjige (301), platforma na koju se knjiga postavlja (302), infracrvena kamera i senzori (315 i 320), projektor (325) i HD kamere (305 i 310). U nekim slučajevima koristi se samo jedna kamera, dok se u drugim može koristiti i više od dvije, ovisno o specifičnostima građe koja se skenira.



Slika 2. Google-ova tehnologija za skeniranje knjiga (Patent #7508978)

Iako Google svakako teži tome, cjelokupni proces nije u potpunosti automatiziran. Prilikom proučavanja tehnologije, utvrđeno je da stanje tadašnje robotike ne dopušta ubrzani proces u kojem bi stroj mogao okretati stranice bez da ih uništi. Dakle, bez obzira na činjenicu da zapošljavanje ogromnog broja ljudi nije bilo u interesu Googlea, ipak su to učinili zbog potrebe ljudske intervencije prilikom okretanja stranica. Što se tiče cijene skeniranja jedne knjige, zahvaljujući velikom broju osoblja koji okreću stranice na digitalizacijskim stanicama, kao i gotovo automatiziranoj obradi, Google Books knjige skenira za samo 10\$ po komadu.

³¹ Shankland, S. Patent reveals Google's book-scanning advantage. 2009. // CNet. Dostupno na: <http://www.cnet.com/news/patent-reveals-googles-book-scanning-advantage/> (17. kolovoza 2016.).

4.4. Projekt Open Content Alliance

Ranije spomenute tužbe, koje su na neki način osporile zakonitost Googleova projekta digitalizacije, osigurale su svojevrsnu odgodu procesa skeniranja, što je naposljetku još jednoj popularnoj tražilici, konkurentu Yahoo!-u omogućilo da se i sam uključi u projekt digitalizacije.³² Godine 2005. Yahoo! je najavio suradnju mnogobrojnih međunarodnih kulturnih, tehnoloških, neprofitnih i nevladinih organizacija koje su počele raditi na novom masovnom projektu digitalizacije s ciljem uspostavljanja fleksibilne, otvorene infrastrukture u kojoj će predstaviti veliku zbirku digitalizirane građe.³³ Projekt je nazvan Open Content Alliance (OCA; Savez otvorenog sadržaja), a temeljni mu je cilj stvaranje digitalne zbirke otvorenog pristupa.

Dakle, OCA je konzorcij ili savez akademskih i narodnih knjižnica i organizacija koje doprinose trajnom i javno dostupnom arhivu digitalnih tekstova. Oni uključuju Boston Library Consortium, California Digital Library, CARLI i mnoge pojedinačne knjižnice kao što je na primjer Britanska knjižnica, ali i one koje surađuju s Googleom – sveučilišne knjižnice Harvard i Illinois. Organizacije poput Adobe Systems Incorporated, HP Labs i O'Reilly Media također su suradnici na projektu. Jedan od zanimljivijih partnera je Microsoft koji se projektu pridružio na samom početku djelovanja preko vlastitog digitalizacijskog projekta Live Book Search.³⁴

Da bi spriječili pravne prepreke s kojima se Google susreće, Savez je odlučio digitalizirati samo onu građu koja je u javnoj domeni, tj. koja nije zaštićena autorskim pravima, te onu za koju su dobili odobrenje od vlasnika autorskih prava.³⁵ Na predstavljanju projekta, Brewster Kahle (glavni pokretač projekta) izjavio je da OCA cilja na 80% knjiga objavljenih između 1923. i 1964. godine koje nisu zaštićene autorskim pravima,³⁶ zatim proširiti polje interesa na tzv. *djela siročad* – kod kojih se autor i izdavač ne mogu pronaći, nadalje na djela koja se više

³² Dye, J. The digital rights issues : behind book digitization projects. // EContent 29, 1(2006). Dostupno na: <http://www.econtentmag.com/Articles/Editorial/Feature/Scanning-the-Stacks-The-Digital-Rights-Issues-Behind-Book-Digitization-Projects-14884.htm> (14. kolovoza 2016.).

³³ Quint, B. . Open Content Alliance expands rapidly. // Information Today 22, 10(2005), str. 7-8. Dostupno na: <http://newsbreaks.infotoday.com/NewsBreaks/Open-Content-Alliance-Expands-Rapidly-Reveals-Operational-Details-16091.asp> (14. kolovoza 2016.).

³⁴ Dougan, K. Music to our eyes: Google Books, Google Scholar, and the Open Content Alliance. // Portal: Libraries and the Academy 10, 1(2010), str. 75-93. Dostupno na: go.gl/zPZryF (15. kolovoza 2016.).

³⁵ Balas, J. L. By digitizing, are we trading future accessibility for current availability? // Computers in Libraries 27, 3(2007), str. 30-32. Dostupno na: <http://connection.ebscohost.com/c/editorials/24353509/by-digitizing-are-we-trading-future-accessibility-current-availability> (13. kolovoza 2016.).

³⁶ Tennant, R. The Open Content Alliance. // Library Journal 130, 20(2005), str. 38. Dostupno na: <http://web.a.ebscohost.com/ehost/pdfviewer/pdfviewer?sid=5da80c72-4879-4274-8768-ce8e8c94d0d3%40sessionmgr4002&vid=0&hid=4201> (18. kolovoza 2016.).

ne tiskaju, te konačno na tiskanu građu. Za distribuciju građe zaštićene autorskim pravima, Savez koristi *Creative Commons* licence koje nude brojne modele licenciranja potičući osobnu upotrebu, ponovnu upotrebu i fleksibilan pristup digitalnim sadržajima.³⁷

Zbirke koje se nalaze u repozitoriju OCA pokrivaju širok raspon građe – uključujući digitaliziranu tiskanu i multimedijalnu građu, a može im se pristupiti preko web stranice Internet Archivea. Osim toga, kopije zbirki pohranjene su i u drugim arhivima u inozemstvu kako bi se osiguralo njihovo dugoročno očuvanje i dostupnost svima.³⁸

Kako bi ju korisnici mogli pronaći i lakše joj pristupiti, Yahoo! indeksira sadržaj sve građe u svojoj tražilici. U međuvremenu, pravila o korištenju i pristupu pojedinim jedinicama građe razlikuju se ovisno o parametrima koje su odredile sudjelujuće institucije. Na primjer, Zbirka američke književnosti koja je produkt suradnje Internet Archivea, Sveučilišta u Kaliforniji i Yahoo!-a, nema nikakva ograničenja, može se preuzeti i koristiti u bilo koju svrhu. Čak postoji i interes među izdavačima za mogućnost predstavljanja radova kroz rad OCA. Na primjer, O'Reilly Media, jedan od osnivačkih članova Saveza koji je tada bio relativno novi izdavač, pristao je besplatno dati određenu građu na raspolaganje OCA u nadi da će postati *vidljivija* korisnicima.

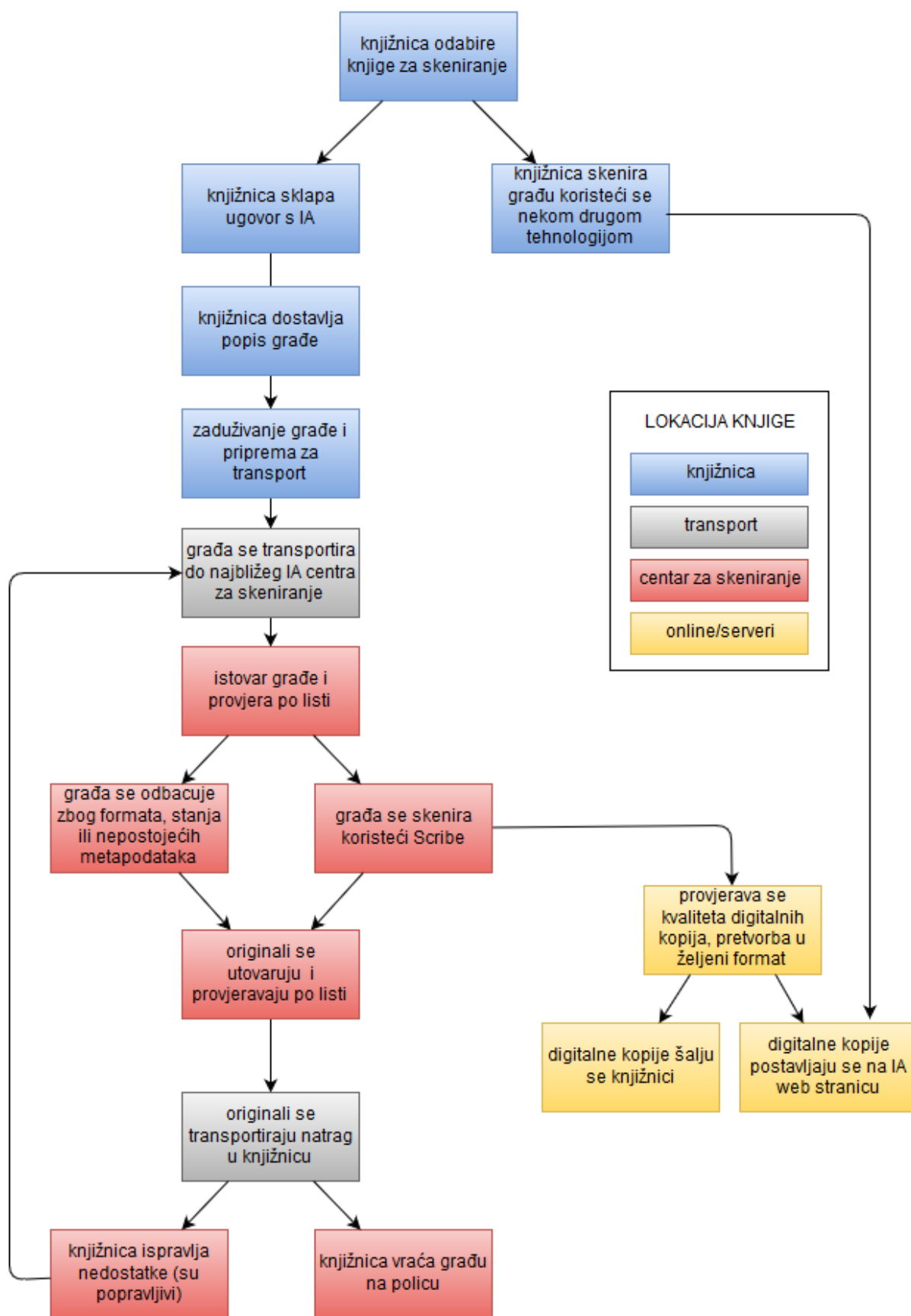
4.4.1. Proces digitalizacije i tehnologija koja se koristi

Brewster Khale, osnivač OCA i Internet Archivea, objašnjavajući doprinose partnera, navodi: „Internet Archive pohranjivat će građu i ponekad pomagati s digitalizacijom, Yahoo će indeksirati sadržaj, Adobe i HP pomažu s programom za obradu...”³⁹ Sam proces digitaliziranja (ilustriran na Slici 3.) vrlo je sličan kao i kod Google Books digitaliziranja. Knjižnice svoje zbirke stavljaju na raspolaganje Savezu, a korporativni sponzori ili Internet Archive osiguravaju sredstva za digitalizaciju. Svaka knjižnica radi s OCA na digitalizacijskom planu, odabirući pritom knjige koje će ući u zbirku za digitalizaciju. Savez osigurava osoblje i opremu za skeniranje. Knjižnično osoblje dostavlja građu u OCA objekt, gdje se ona skenira, te se po završetku vraća u knjižnicu.

³⁷ Quint, B. Open Content Alliance rises to the challenge of Google Print. 3. listopada 2005. // Information Today. Dostupno na: <http://newsbreaks.infoday.com/NewsBreaks/Open-Content-Alliance-Rises-to-the-Challenge-of-Google-Print-16110.asp> (15. kolovoza 2016.).

³⁸ www.opencontentalliance.org

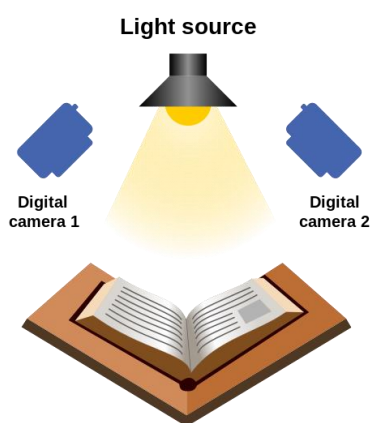
³⁹ Albanese, A. Yahoo, partners debut scan plan. // Library Journal 130, 17(2005), str. 14-15. Dostupno na: <http://edc-connection.ebscohost.com/c/reports/18576033/yahoo-partners-debut-scan-plan> (15. kolovoza 2016.).



Slika 3. Internet Archive – tijekom digitalizacije knjiga (prema Jones, 2014)

Za razliku od Google Books-ove početne tajnovitosti o vlastitoj digitalizacijskoj opremi,⁴⁰ Open Content Alliance objavio je opsežne detalje o svom Scribe sustavu, kao i o drugim opcijama za suradnike i korisnike. Unatoč početnom istraživanju robotizirane tehnologije koja bi sama okretala stranice knjige, Savez se ipak odlučio na model u kojem se stranice ručno okreću i razvio prilagođeni sustav skeniranja pod nazivom Scribe. Radna stanica sastoji se od čvrstog aluminijskog okvira koji drži dvije podesive vodilice za kamere, dvije kamere koje slikaju recto i verso stranice svake knjige, platformu za knjigu u obliku slova V, staklenu ploču istog oblika koja se podiže i spušta pomoću pedale, dva rasvjetna tijela koja osvijetljavaju knjigu te kompjuter na koji se preuzimaju i obrađuju slike.⁴¹ Operater koji radi na Scribe-u, okreće stranice knjige i pri svakom okretaju pritišće pedalu koja na knjigu spušta staklenu ploču kako bi se stranice poravnale za fotografiranje. Kvaliteta svake skenirane stranice dodatno se provjerava i po potrebi se ručno podešava. Prosječni radnik u sat vremena skenira oko 350 stranica (u prosjeku jedna stranica svakih 10 sekundi).

Dok Google koristi ravnu površinu, oslanjajući se prvenstveno na softver i algoritme za izravnavanje i poboljšanje slike, sustav Internet Archivea snažnije se oslanja na hardver kako bi osigurao kvalitetu – knjiga se utisne u platformu u obliku slova V kako bi se poravnale stranice, što odmah smanjuje zakrivljenost. Rezultat toga je čišća slika koja dolazi ravno iz skenera. Zbog toga, za razliku od Googlea, nije potrebno puno ulagati u naknadnu obradu slike. Međutim, vrlo dobra slika koja dolazi ravno iz skenera ipak dolazi pod cijenom brzine – budući da je potrebno više vremena za podizanje i spuštanje stakla natrag na knjigu, cjelokupni proces u osnovi je sporiji od Googleovog.



Slika 4. Ilustracija tehnologije koju koristi OCA (lijevo) i fotografija Scribe jedinice za skeniranje (desno)

⁴⁰ Ashmore, B., Grogg, J. E. The race to the shelf continues : the Open Content Alliance. // Searcher 16, 1 (2008), str. 18-23. Dostupno na: http://www.infoday.com/searcher/jan08/Ashmore_Grogg.shtml (16. kolovoza 2016.).

⁴¹ Jones, E. Constructing the universal library : doktorski rad. Washington : University of Washington, 2014.

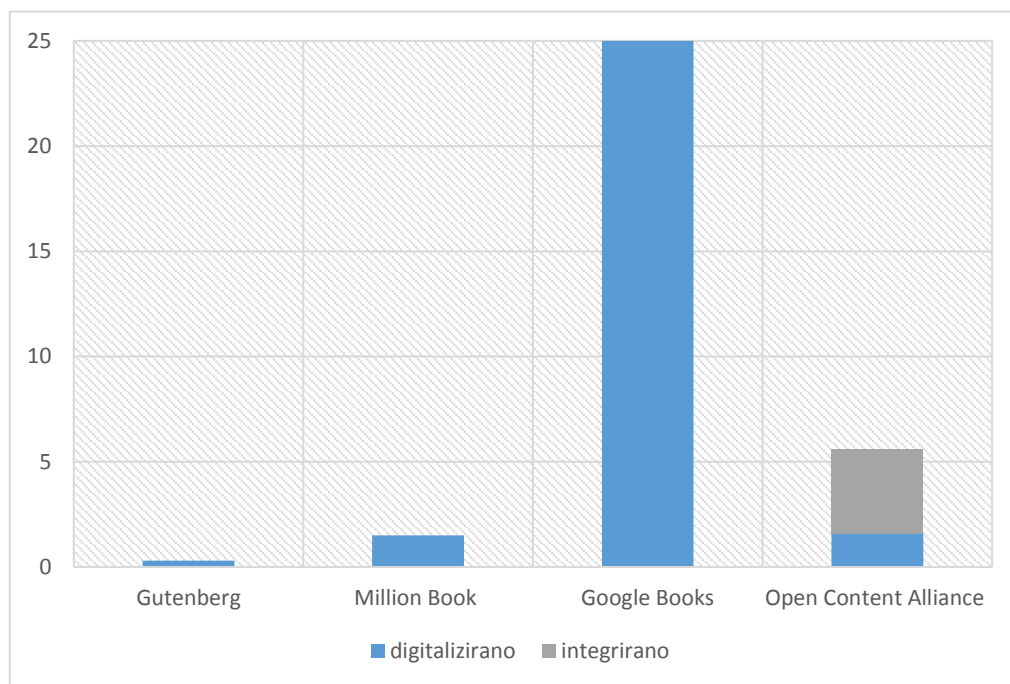
Od 2005. do 2009. godine Open Content Alliance proizveo je dosta veliku količinu digitaliziranih knjiga – više od 1 milijun, zagovarajući pritom otvorenost pristupa. A onda, sredinom 2009. godine, kako navodi Jones (2014) „it quietly collapsed“. Bez Microsoftove financijske potpore i vlastitog čvrstog vodstva, Savez je tiho propao – ili barem izgubio koheziju koju je jednom posjedovao. OCA, odnosno suradnici koji su do tada radili na projektu, podijelili su se u nekoliko svojevrsnih projekata nasljednika – digitalno posuđivanje i međuknjižnična posudba, servisi skeniranja na zahtjev, arhivi fizičkih knjiga i još mnogi drugi. U mnogima od njih, prijelaz s OCA na druge projekte bio je gotovo neprimjetan, a pojam Open Content Alliance polako je izlazio iz upotrebe, iako su se s njim povezane djelatnosti i aktivnosti – nastavile.

Internet Archive i dalje pruža pristup digitaliziranoj građi OCA kroz svoju *Ebook and Texts* arhivu te *Open Library* portal u kojima su dostupni i milijuni novih digitaliziranih knjiga (više od 10 milijuna⁴²). Osim kroz rad Internet Archive, moglo bi se reći da OCA i dalje živi u mnogim drugim projektima poput Digital Public Library of America i lokalnih inicijativa. Zbirka skeniranih djela OCA pomogla je i u stvaranju HathiTrust-a koji je osnovan kao kolaborativna inicijativa između knjižnica i partnera koji su većinom surađivali i u OCA.

⁴² eBooks and Texts. // Internet Archive. Dostupno na: <https://archive.org/details/texts> (20. kolovoza 2016.).

5. Usporedna analiza projekata Open Content Alliance i Google Books

Da nema drugih zanimljivih značajki, projekt Google Books isticao bi se i samo na temelju količine digitalizirane građe. Broj svezaka koje je skenirao, tempo i brzina kojim ih je skenirao, nalaze se na potpuno drugačijoj razini od ostalih projekata. Najstariji projekt – projekt Gutenberg, trenutno uključuje oko 100 tisuća svezaka koje su od 1971. godine pretipkavali, a zatim i skenirali volonteri koji su radili na projektu. The Million Book skenirao je 1,5 milijun svezaka između 2000. i 2008. godine.⁴³ Projekt Open Content Alliance sveukupno je skenirao (1,6mil) i integrirao relativno impresivnih 5,6 milijuna svezaka u 8 godina. Međutim, Google Books nadmašio je i tu brojku. U 12 godina (2004.-2015.) Google je skenirao nešto više od 25 milijuna knjiga,⁴⁴ na temelju čega se može zaključiti da njegova prosječna stopa digitalizacije iznosi nešto više od 2 milijuna knjiga godišnje – gotovo pola brojke cjelokupne digitalizirane i integrirane građe OCA. Slika 2. ilustrira golemost razlike u broju skenirane građe između Google Books-a i ostala tri spomenuta projekta.



Slika 5. Broj skeniranih knjiga (u milijunima) do kraja 2015. godine

⁴³ St. Clair, G. The Million Book project in relation to Google. // Journal of Library Administration 47, 1/2(2008), str. 151-163. <http://www.tandfonline.com/doi/abs/10.1080/01930820802111041> (17. kolovoza 2016.).

⁴⁴ Heyman, S. Google Books : a complex and controversial experiment. 2015. // The New York Times Magazine. Dostupno na: http://www.nytimes.com/2015/10/29/arts/international/google-books-a-complex-and-controversial-experiment.html?_r=0 (17. kolovoza 2016.).

Djelomično kao rezultat digitalizacije u masovnim razmjerima, Google Books je među članovima šire javnosti postao daleko najpoznatiji projekt svoje vrste,⁴⁵ a kao takav vjerojatno i najkorišteniji. Nadalje, njegov uspjeh dokaz je održivosti privatnog, centraliziranog modela digitalizacije, a spor i kontroverze oko autorskih prava, Googleov tim i partnerske knjižnice stavio je u jedinstvenu poziciju s koje mogu utjecati ne samo na oblikovanje daljnjeg razvoja digitalizacije, već i na oblikovanje digitalnih autorskih prava. Ti čimbenici u kombinaciji sa sveprisutnom Googleovim utjecajem u svijetu, ukazuju na to da ono što Google sada radi, imat će snažan utjecaj na buduće strategije digitalizacije kao i sustava u kojem će se razvijati – bilo da će ga buduće inicijative oponašati ili se drastično razlikovati od njega.

Open Content Alliance, iako znatno manji, ima smisla analizirati ga usporedno s Google Books-om zbog nekoliko ključnih razloga. Prvi i možda najvažniji, OCA je osnovan poprilično eksplicitno kao kontrapunkt Google Books-u.⁴⁶ Njegova glavna koordinacijska organizacija – Internet Archive, knjige je skenirala godinama prije nego što je OCA osnovan. Međutim, Googleova najava kao da je potakla osnivača Internet Archivea – Brewstera Kahlea, da pokrene alternativu tom projektu koja će biti otvoreniji i kolaborativniji model od Googleovog. To je i učinjeno. Za razliku od Googleovog projekta, OCA je mnogo više decentraliziran – i mehanički i administrativno. U velikoj se mjeri oslanja na to da digitalizaciju provodi u suradnji s partnerima i ne preuzima centraliziranu kontrolu nad resursima koje zajedno prikupljaju. Iako je Savez imao vlastite centre za skeniranje, u svoju je online zbirku integrirao i građu digitaliziranu u drugim projektima poput Million Book (100,000+) i Gutenberg (36,000+).⁴⁷ Nadalje, OCA se snažnije od Googlea usmjerava na visoko kvalitetno skeniranje čiji se proizvodi, tj. kopije visoke razlučivosti direktno daju na raspolaganje korisnicima, bez potrebe za dodatnom obradom samih slika.

⁴⁵ St. Clair, G. The Million Book project in relation to Google. // *Journal of Library Administration* 47, 1/2(2008), str. 151-163. Dostupno na: <http://www.tandfonline.com/doi/abs/10.1080/01930820802111041> (17. kolovoza 2016.).

⁴⁶ Albanese, A. R. Scan this book! // *Library Journal* 132, 13(2007), str. 32-35. Dostupno na: <http://lj.libraryjournal.com/2007/08/technology/scan-this-book/> (14. kolovoza 2016.).

⁴⁷ Jones, E. *Constructing the universal library* : doktorski rad. Washington : University of Washington, 2014.

Tablica 1. Ciljevi i karakteristične značajke projekata Google Books i Open Content Alliance

Google Books	Open Content Alliance
<ul style="list-style-type: none"> • pružiti sveobuhvatan, pretraživ indeks objavljenih knjiga na različitim jezicima • korisnicima olakšati pretraživanje i otkrivanje relevantnih knjiga preko Googleove tražilice • privući što veći broj korisnika • digitalizirati svaku knjigu ikad objavljenu kako bi ih korisnici lakše otkrili 	<ul style="list-style-type: none"> • stvoriti digitalnu zbirku otvorenog pristupa • podržati razvoj trajnog arhiva digitaliziranih višejezičnih tekstova i multimedijских sadržaja • očuvati datoteke pohranjujući ih u više repozitorija • stvoriti trajnu arhivu znanstvenih i akademskih djela do čijeg bi se sadržaja moglo doći pretraživanjem bilo koje tražilice

Iako su Google Books i Open Content Alliance naizgled veoma različite inicijative, ipak imaju dosta toga zajedničkog i dijele mnoge sličnosti. Kao prvo, obje inicijative financijsku pomoć i tehnološku stručnost crpe, ili su crpili, iz velikih korporacija. Dok je Google Books financiran i upravljan od strane diva kao što je Google, Savezu se u samom početku, osim Yahoo!-a, pridružio i Microsoft s otprilike 5 milijuna dolara pomoći i obećanjem da će digitalizirati oko 150 tisuća knjiga.⁴⁸ Međutim, Microsoft je naknadno povukao sudjelovanje u Savezu.⁴⁹

Iduća sličnost između dvije inicijative njihova je arhivistička funkcija. Kroz procese digitalizacije, oba projekta doprinose stvaranju trajne digitalne zbirke građe koje se mogu koristiti kao kopije krhkih i raspadajućih dokumenata. Osim toga mogu biti i od neprocjenjive vrijednosti u slučajevima gdje poplave, požari i ostale katastrofe oštete ili unište originalnu građu. Čak i da se nešto dogodi s Googleom ili Internet Archiveom, da postanu nedostupni ili nešto drugo, kopije se čuvaju u repozitorijima knjižnica koje su surađivale na projektu – što nas dovodi do iduće sličnosti. Oba projekta digitaliziranu građu dijeli s knjižnicama od kojih su došli originali, te im dopuštaju daljnju distribuciju elektroničkih inačica unutar vlastite mreže korisnika.

⁴⁸ Quint, B. . Open Content Alliance expands rapidly. 31. listopada 2005. // Information Today. Dostupno na: <http://newsbreaks.infoday.com/NewsBreaks/Open-Content-Alliance-Expands-Rapidly-Reveals-Operational-Details-16091.asp> (14. kolovoza 2016.).

⁴⁹ Ashmore, B., Grogg, J. E. The race to the shelf continues : the Open Content Alliance. // Searcher 16, 1 (2008), str. 18-23. Dostupno na: http://www.infoday.com/searcher/jan08/Ashmore_Grogg.shtml (16. kolovoza 2016.).

Nadalje, posluju kao svojevrsni *outsourcing* agenti. Knjižnica koja sudjeluje u projektu, bilo da se radi o OCA ili Google Books, prevozi knjige do objekta u kojem se skeniraju i vraća ih kada je postupak gotov. Knjižničnom osoblju upravljanje jedinicama za digitalizaciju nije dozvoljeno. Također, oba projekta, iako se tehnologije razlikuju, koriste model ručnog okretanja stranica i imaju velike timove osoblja koji upravljaju jedinicama.

Iako je u osnovama oba projekta obavljanje digitalizacije i stavljanje na raspolaganje korisnicima digitalizirane tekstove, postoji niz ključnih razlika koje je važno spomenuti. Te se razlike ponajprije nalaze u načinu na koji interpretiraju autorska prava, u kojoj su mjeri spremni podijeliti svoje metode i tehnologije u javnosti, u njihovim motivima za digitaliziranje, te na kraju u potrebama krajnjih korisnika koje namjeravaju ispuniti. Na primjer, potrebe korisnika koji je u potrazi za prijedlozima koju knjigu pročitati/kupiti, vrlo su različite od potreba istraživača i povjesničara koji su zainteresirani za pažljivo čitanje i proučavanje cjelokupnog sadržaja povijesno značajnih knjiga. Dok su funkcije Google Books-a dovoljne da bi ispunile potrebe prvog korisnika, drugi korisnik trebao bi potpun i nesmetan pristup knjizi što Google ne može uvijek omogućiti.

Projekti se također razlikuju i po pitanju problematike autorskih prava – OCA se odlučio za potpuno drugačije rješenje. Google prvo sve skenira ne obazirući se pritom na status autorskih prava, a tek nakon digitaliziranja nudi različite razine pristupa ovisno o statusu (ipak, vlasnicima autorskih prava ponuđena je opcija *opt-out*, kojom u stvari stavljaju svojevrsnu zabranu na prikazivanje knjige). S druge strane, strategija Open Content Alliancea, tj. Internet Archivea je ta da prije samog procesa digitalizacije utvrde da je djelo u javnoj domeni, odnosno da traže dopuštenje za skeniranje od vlasnika autorskih prava. Ultimativni cilj obje inicijative bio je u konačnici digitalizirati svaku knjigu ikad objavljenu. Međutim, kao rezultat zakona o autorskim pravima kojih se držao OCA, u kombinaciji s preprekama u administraciji i financiranju, provodio se puno drugačiji red i tempo digitalizacije, kao što je već spomenuto.

Jedno od važnijih pitanja vezanih uz problematiku Google Books-a upravo je pitanje pripadaju li njegove aktivnosti u iznimke od autorskih prava (eng. *fair-use*). Google je tvrdio upravo to pod opravdanjem da se digitalizirana djela slobodno mogu koristiti u određene svrhe (npr. citiranje ili nastava). Unatoč mnogim kritičarima te autorima i izdavačima koji su se oštro suprotstavljali projektu, u travnju 2016. godine, nakon desetogodišnje sudske bitke između Autorske komore (kao i Udruge američkih nakladnika) i projekta Google Books, Vrhovni sud SAD-a odbio je čuti još jednu u nizu žalbi Autorske komore. Tim je činom

Vrhovni sud potvrdio odluku nižeg suda da je Googleu dopušteno skeniranje knjiga i prikaz isječaka u rezultatima pretraživanja, što bi značilo da projekt ne krši zakon niti autorska prava.⁵⁰

Nadalje, vezano uz razlike između dva projekta, Quint (2005) navodi da je, za razliku od GB projekta, OCA puno otvoreniji o tehnologiji i opremi koju koristi, kao i u dijeljenju proizvoda procesa digitalizacije. Prema Balas (2007), upravo je činjenica da je Google želio da se rezultati pretraživanja Google Books-a pojavljuju samo u njegovoj tražilici (a ne npr. i u Bing, Yahoo! i ostalima), ta koja je potakla stvaranje OCA. Johnson (2007) napominje da sve knjižnice koje se žele uključiti u Google Books moraju pristati na niz uvjeta između ostalog, uključuju obvezu da digitaliziranu građu učine nedostupnom drugim tražilicama. Nasuprot tome, sva građa OCA dostupna je na bilo kojoj usluzi pretraživanja, uključujući i Google.⁵¹

Kada je riječ o motivima za digitalizaciju, Google, komercijalni, s profitabilnošću i dobiti upravljan div, očito ima koristi od uključivanja projekta digitalizacije u svoj arsenal usluga. Prihod je toliko značajan da Google u stvari plaća knjižnicama da bi dobio pristup njihovim knjigama.⁵² OCA, s druge strane, nema monetaran dobitak, a trošak skeniranja dijele članovi Saveza (30\$ po knjizi).⁵³ Prema tome, očito je da Google Books posjeduje određene financijske prednosti za knjižnice koje se odluče priključiti projektu, budući da su knjižnice koje se priključe Savezu dužne platiti troškove skeniranja.

Nadovezujući se na financijsku stranu problematike, dugoročna dostupnost bilo koje digitalne zbirke prije svega ovisi o sredstvima koja ju održavaju. U slučaju projekta Google Books, uz digitaliziranu građu često se pojavljuju i reklame. Unatoč tome što to donekle *komercijalizira* korisničko iskustvo, ipak daje određene prihode što projekt čini samodostatnim i osigurava dugoročnu financijsku održivost. OCA, s druge strane, isključivo ovisi o doprinosima osnivača i korporativnih sponzora, a i od knjižnica zahtijeva da subvencioniraju barem dio procesa. Upravo zbog gubitka većih sponzora, a i uz sukobe oko toga kako bi se trebao voditi, Open Content Alliance postupno je prestao postojati.

⁵⁰ Heyman, S. Google Books : a complex and controversial experiment. 2015. // The New York Times Magazine. Dostupno na: http://www.nytimes.com/2015/10/29/arts/international/google-books-a-complex-and-controversial-experiment.html?_r=0 (17. kolovoza 2016.).

⁵¹ Hafner, K. Libraries rebuff Google and Microsoft on offers to place books on web. 22. listopada 2007. // The New York Times Magazine. Dostupno na: <http://www.nytimes.com/2007/10/22/technology/22iht-library.1.7992196.html> (15. kolovoza 2016.).

⁵² Ibid.

⁵³ Tennant, R. The Open Content Alliance. // Library Journal 130, 20(2005), str. 38. Dostupno na: <http://web.a.ebscohost.com/ehost/pdfviewer/pdfviewer?sid=5da80c72-4879-4274-8768-ce8e8c94d0d3%40sessionmgr4002&vid=0&hid=4201> (18. kolovoza 2016.).

Zaključak

Postojanje inicijativa poput Google Books i Open Content Alliance pogoduje knjižnicama i njihovim korisnicima, čak i kada knjižnice ne sudjeluju u njima. Budući da je gotovo cjeloviti tekst svakodnevno rastućeg broja digitaliziranih knjiga besplatno dostupan online, putem Googlea ili Internet Archive-a, korisnici u potrazi za informacijom ne moraju više prelistavati knjige ili koristiti kataloge da bi ju pronašli. Umjesto toga, mogu koristiti pretraživače koji će im traženu informaciju prikazati u kontekstu. To ne samo da rezultira uštedom vremena, već i smanjenjem oštećenja na rijetkim i krhkim knjigama. Nadalje, oni korisnici koji nemaju pristup knjižnici ili pak rijetkim i povijesno značajnim knjigama, te knjige sada mogu čitati – u cijelosti ili djelomično, i to besplatno putem Interneta.

Jasno je da ove dvije inicijative, i druge poput njih, mijenjaju način na koji ljudi pristupaju informacijama – na knjižničarima je da odluče hoće li se priključiti pokretu digitalizacije, te u kojoj mjeri će to učiniti. Svijest o postojanju digitalnih kopija vrijednih knjiga malen je korak prema sve većem uvođenju e-knjiga u knjižnične zbirke – a možda i jedina promjena koju će neke knjižnice napraviti. Mnoge su knjižnice zasigurno već počele sa svojim malim projektima digitalizacije kako bi očuvali krhke ili posebno vrijedne knjige, ili pak poboljšale pristup važnoj građi udaljenijim korisnicima. Međutim, kao što se može i zaključiti, budući da je trošak osoblja i opreme potrebne za digitalizaciju poprilično velik, korisno je za knjižnice da postanu dio nekog većeg projekta gdje se sredstva i trošak mogu dijeliti. Ipak, prije nego što će se pridružiti takvim inicijativama, vrlo je važno raspraviti o tome kako će se postupati s autorskim pravima njihove građe, s kojim će se ograničenjima upotrebe susresti, ali i kakvu korist (novčanu ili neku drugu) očekuju od postavljanja svojih knjiga online.

Nadalje, često se postavljaju pitanja o budućnosti knjižnica i načinu na koji ljudi koriste informacije. Činjenica je da je digitalizacija itekako promijenila koncept knjižnica i način na koji korisnici, ne samo koriste, već i pristupaju informacijama. Unatoč mračnim prognozama da to označava *početak kraja* knjižnica, mnogi ipak projekte digitalizacije smatraju značajnim za razvoj cjelokupnog društva,⁵⁴ a osobito za knjižnice i bibliotekarstvo kao struku koja ima priliku dobiti još dublji smisao. Ustaljena je pojava da se dolaskom svake nove tehnologije i medija najavljuje i svojevrsni kraj knjižnica, pa su slične najave pratile i širenje Interneta.

⁵⁴ Šapro-Ficović, M. Masovna digitalizacija knjiga : utjecaj na knjižnice. // Vjesnik bibliotekara Hrvatske. 54, 1/2(2011), str. 216-250. Dostupno na: <http://www.hkdrustvo.hr/vbh/broj/103> (15. kolovoza 2016.).

Povezano s time, Gorman (2006) zaključuje da današnje vrijeme i položaj u kojem se knjižnice nalaze nije ni po čemu poseban u odnosu na prošla vremena – logičnije je pretpostaviti da se nalazimo na određenoj točki ravnomjerne evolucije knjižnica i ljudske komunikacije, u kojoj je jednako toliko promjena iza nas koliko i pred nama. Takvim razumijevanjem stvari, autor napominje da se kao struka možemo opustiti i hladnokrvno rješavati svoje probleme primjenjujući analizu i logiku – *evolucijska vremena zahtijevaju evolucijske odgovore*.⁵⁵

Budući da su knjižnice prije svega društvene ustanove s ciljem poticanja informiranja, obrazovanja, rekreativnih aktivnosti i osobnog rasta pojedinca te naposljetku boljitka zajednice, nove tehnologije i mediji ne mogu znatnije ugroziti njihov opstanak. Ono što im zapravo omogućavaju su prilike za značajne i svrhovite promjene i preuzimanje novih zadaća.

Prema tome, nema ništa lošeg u tome što je Google, kao ogromna korporacija, u polje svojih interesa uvrstio i ovu, naočigled bibliotekarsku preokupaciju digitaliziranja knjiga. Da se Google nije upustio u svoj projekt i pokazao svijetu da je masovna digitalizacija knjižnične građe i knjižničnih zbirki uistinu moguća, moguće je pretpostaviti da bi korpus digitaliziranih djela brojio nekoliko desetaka tisuća, umjesto milijuna djela. Ipak, ne možemo znati što bi se dogodilo da Google nije započeo sa svojim projektom digitalizacije. Ono što je sigurno jest da u to vrijeme to nitko nije radio. Nitko to ne radi ni danas.

⁵⁵ Gorman, M. Postojana knjižnica : tehnologija, tradicija i potraga za ravnotežom. Zagreb : Hrvatsko knjižničarsko društvo, 2006.

Literatura

1. Albanese, A. R. Scan this book! // Library Journal 132, 13(2007), str. 32-35. Dostupno na: <http://lj.libraryjournal.com/2007/08/technology/scan-this-book/> (14. kolovoza 2016.).
2. Albanese, A. Yahoo, partners debut scan plan. // Library Journal 130, 17(2005), str. 14-15. Dostupno na: <http://edc-connection.ebscohost.com/c/reports/18576033/yahoo-partners-debut-scan-plan> (15. kolovoza 2016.).
3. Ashmore, B., Grogg, J. E. The race to the shelf continues : the Open Content Alliance. // Searcher 16, 1 (2008), str. 18-23. Dostupno na: http://www.infoday.com/searcher/jan08/Ashmore_Grogg.shtml (16. kolovoza 2016.).
4. Balas, J. L. By digitizing, are we trading future accessibility for current availability? // Computers in Libraries 27, 3(2007), str. 30-32. Dostupno na: <http://connection.ebscohost.com/c/editorials/24353509/by-digitizing-are-we-trading-future-accessibility-current-availability> (13. kolovoza 2016.).
5. Clements, M. The secret of Google's book scanning machine revealed. 2009. // National Public Radio. Dostupno na: http://www.npr.org/sections/library/2009/04/the_granting_of_patent_7508978.html (17. kolovoza 2016.).
6. Coyle, K. Mass digitization of books. // Journal of Academic Librarianship 32, 6(2006). Dostupno na: <http://www.kcoyle.net/jal-32-6.html> (15. kolovoza 2016.).
7. Deegan, M., Sutherland, K. Transferred illusions : digital technology and the forms of print. Farnham: Ashgate, 2009.
8. Digitalizacija. // Hrvatska enciklopedija. Mrežno izd. Zagreb : Leksikografski zavod Miroslav Krleža. Dostupno na: <http://www.enciklopedija.hr/natuknica.aspx?id=68025> (13. kolovoza 2016.).
9. Dougan, K. Music to our eyes : Google Books, Google Scholar, and the Open Content Alliance. // Portal: Libraries and the Academy 10, 1(2010), str. 75-93. Dostupno na: goo.gl/zPZryF (15. kolovoza 2016.).

10. Dye, J. The digital rights issues : behind book digitization projects. // EContent 29, 1(2006). Dostupno na:
<http://www.econtentmag.com/Articles/Editorial/Feature/Scanning-the-Stacks-The-Digital-Rights-Issues-Behind-Book-Digitization-Projects-14884.htm> (14. kolovoza 2016.).
11. Free ebooks by project Gutenberg. // Project Gutenberg. Dostupno na:
http://www.gutenberg.org/wiki/Main_Page (15. kolovoza 2016.).
12. Gorman, M. Postojana knjižnica : tehnologija, tradicija i potraga za ravnotežom. Zagreb : Hrvatsko knjižničarsko društvo, 2006.
13. Hafner, K. Libraries rebuff Google and Microsoft on offers to place books on web. 22. listopada 2007. // The New York Times Magazine. Dostupno na:
<http://www.nytimes.com/2007/10/22/technology/22iht-library.1.7992196.html> (15.kolovoza 2016.).
14. Hart, M. Project Gutenberg mission statement. 2007. // Project Gutenberg. Dostupno na:
http://www.gutenberg.org/wiki/Gutenberg:Project_Gutenberg_Mission_Statement_by_Michael_Hart (15. kolovoza 2016.).
15. Hart, M. The history and philosophy of project Gutenberg. 1992. // Project Gutenberg. Dostupno na:
http://www.gutenberg.org/wiki/Gutenberg:The_History_and_Philosophy_of_Project_Gutenberg_by_Michael_Hart (15. kolovoza 2016.).
16. Heyman, S. Google Books : a complex and controversial experiment. 2015. // The New York Times Magazine. Dostupno na:
http://www.nytimes.com/2015/10/29/arts/international/google-books-a-complex-and-controversial-experiment.html?_r=0 (17. kolovoza 2016.).
17. Horvat, A., Živković, D. Između javnosti i privatnosti : knjižnice u vremenu e-knjige. Zagreb : Hrvatska sveučilišna naklada, 2012.
18. Johnson, R. K. In Google's broad wake : taking responsibility for shaping the global digital library. // Association of Research Libraries 250(2007), str. 1-15. Dostupno na:
<http://www.arl.org/storage/documents/publications/ar1-br-250.pdf> (15. kolovoza 2016.).
19. Jones, E. Constructing the universal library : doktorski rad. Washington : University of Washington, 2014.

20. Kahle, B. Announcing the Open Content Alliance. 2005. // Yahoo! Search Blog.
Dostupno na: <http://www.ysearchblog.com/2005/10/02/announcing-the-open-content-alliance/> (15. kolovoza 2016.).
21. Karle–Zenith, A. Google Book Search and the University of Michigan. 20. veljače 2007. Dostupno na: http://eprints.rclis.org/9011/1/Zenith_Google.pdf (15. kolovoza 2016.).
22. Neomedia Komunikacije. Priprema i planiranje projekta digitalizacije. 2008. // Hrvatska kulturna baština. Dostupno na: <http://www.kultura.hr/Sudjelujte/Digitalizacija/Priprema-i-planiranje-projekta-digitalizacije> (17. kolovoza 2016.).
23. O'Leary, M. Open Content Alliance embodies open source movement. // Information Today 26, 1(2009), str. 37-43. Dostupno na: <http://web.a.ebscohost.com/ehost/pdfviewer/pdfviewer?sid=707934aa-f446-4abb-985c-a2835f27e536%40sessionmgr4005&vid=0&hid=4201> (17. kolovoza 2016.).
24. Online library gives readers access to 1.5 million books. 2007. // Carnegie Mellon University. Dostupno na: https://www.cmu.edu/news/archive/2007/November/nov27_ulib.shtml (15. kolovoza 2016.).
25. Quint, B. Open Content Alliance expands rapidly. 31. listopada 2005. // Information Today. Dostupno na: <http://newsbreaks.infotoday.com/NewsBreaks/Open-Content-Alliance-Expands-Rapidly-Reveals-Operational-Details-16091.asp> (14. kolovoza 2016.).
26. Quint, B. Open Content Alliance rises to the challenge of Google Print. 3. listopada 2005. // Information Today. Dostupno na: <http://newsbreaks.infotoday.com/NewsBreaks/Open-Content-Alliance-Rises-to-the-Challenge-of-Google-Print-16110.asp> (15. kolovoza 2016.).
27. Rieger, O. Preservation in the age of large-scale digitization. Washington, D.C.: Council on Library and Information Resources, 2008. Dostupno na: <https://www.clir.org/pubs/reports/pub141/pub141.pdf> (16. kolovoza 2016.).
28. Said, C. Revolutionary chapter : Google's ambitious book-scanning plan seen as key shift in paper-based culture. 20. prosinca 2004. // San Francisco Chronicle. Dostupno na: <http://www.sfgate.com/business/article/Revolutionary-chapter-Google-s-ambitious-2662491.php> (13. kolovoza 2016.).

29. Shankland, S. Patent reveals Google's book-scanning advantage. 2009. // CNet.
Dostupno na: <http://www.cnet.com/news/patent-reveals-googles-book-scanning-advantage/> (17. kolovoza 2016.).
30. Smjernice za odabir građe za digitalizaciju (radna verzija), 2007. // Nacionalni projekt „Hrvatska kulturna baština“, Ministarstvo kulture Republike Hrvatske. Dostupno na: https://bib.irb.hr/datoteka/590089.smjernice_odabir.pdf (14. kolovoza 2016.).
31. St. Clair, G. The Million Book project in relation to Google. // Journal of Library Administration 47, 1/2(2008), str. 151-163. Dostupno na: <http://www.tandfonline.com/doi/abs/10.1080/01930820802111041> (17. kolovoza 2016.).
32. Stančić, H. Digitalizacija. Zagreb : Zavod za informacijske studije Odsjeka za informacijske znanosti Filozofskog fakulteta Sveučilišta u Zagrebu, 2009.
33. Šapro-Ficović, M. Masovna digitalizacija knjiga : utjecaj na knjižnice. // Vjesnik bibliotekara Hrvatske. 54, 1/2(2011), str. 216-250. Dostupno i na: <http://www.hkdrustvo.hr/vbh/broj/103> (15. kolovoza 2016.).
34. Tennant, R. The Open Content Alliance. // Library Journal 130, 20(2005), str. 38.
Dostupno na: <http://web.a.ebscohost.com/ehost/pdfviewer/pdfviewer?sid=5da80c72-4879-4274-8768-ce8e8c94d0d3%40sessionmgr4002&vid=0&hid=4201> (18.kolovoza 2016.).
35. Youngs, K. Managing the digitisation of library, archive and museum materials. London: NPO, 2001. Dostupno na: <http://www.collectionstrust.org.uk/media/documents/c1/a105/f6/000505.pdf> (14. kolovoza 2016.).